PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07H 21/00 — A2

(11) International Publication Number: **WO 98/15567**

(43) International Publication Date: 16 April 1998 (16.04.98)

(21) International Application Number: PCT/IB97/01385

(22) International Filing Date: 6 October 1997 (06.10.97)

(30) Priority Data:
60/027,896 7 October 1996 (07.10.96) US

(71) Applicant (*for all designated States except US*): VITALITY BIOTECHNOLOGIES LTD. [IL/IL]; Einstein Boulevard, Carmel Industrial Park, 39101 Tirat Hacarmel (IL).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): STRIZHOV, Nicolai [RU/DE]; Max-Planck-Institute für Züchtungsforschung, D-5000 Köln (DE). KONCZ, Csaba [HU/DE]; Max-Plank-Institute für Züchtungsforschung, D-5000 Köln (DE). SCHELL, Jeff [DE/DE]; Max-Plank-Institute für Züchtungsforschung, D-5000 Köln (DE).

(74) Agent: OLGEMÖLLER, Luitgard; Leonhard, Olgemöller, Fricke, Josephspitalstrasse 7, D-80331 München (DE).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: GENE SYNTHESIS METHOD

(57) Abstract

A method of gene synthesis is disclosed. The gene synthesis method permits the codons of a natural gene to be changed to allow preferential transcription and translation of the synthetic gene in transgenic organisms. The method utilizes a combination of enzymatic and chemical synthesis of DNA and significantly reduces the cost, time and number of steps required for construction of synthetic genes.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

TITLE: *GENE SYNTHESIS METHOD*

This application is a continuation of U.S. Provisional Application serial number 60/027,896 filed on October 7, 1996 and incorporated herein by reference.

BACKGROUND OF THE INVENTION

Gene Synthesis Techniques

The present invention relates to a method of gene synthesis and more particularly to a gene synthesis method wherein the codons of a natural gene can be changed to allow optimal transcription and translation of the gene in transgenic organisms which prefer particular codons and other sequences that differ from those of the natural gene.

The degeneracy of the genetic code permits substantial freedom in the choice of codons for any particular amino acid sequence. Transgenic organisms such as plants frequently prefer particular codons which, though they encode the same protein, may differ from the codons in the organism from which the gene was derived. For example, U.S. Patent No. 5,380,831 to Adang et al., relates to the creation of insect resistant transgenic plants that express the *Bacillus thuringiensis* (*Bt*) toxin gene. The *Bt* crystal protein, an insect toxin, is encoded by a full-length gene that is poorly expressed in transgenic plants. In order to improve expression in plants, a synthetic gene encoding the protein containing codons preferred in plants was substituted for the natural sequence. The invention disclosed therein comprised a chemically synthesized gene encoding an insecticidal protein which is frequently equivalent to a native insecticidal protein of *Bt*. The synthetic gene was designed to be expressed in plants at a level higher than a native *Bt* gene.

The approach used to assemble the *Bt* gene in the '831 patent is one of several methods for gene synthesis for genetic engineering manipulation. It consisted of designing an improved nucleotide sequence for the coding region and assembling the gene from a number of short chemically synthesized oligonucleotide segments. The DNA sequence to be synthesized was divided into segments, with lengths that can be synthesized, isolated and purified. The segments were then joined enzymatically to form the synthetic *Bt* gene.

Disadvantages of the gene synthesis method described in the '831 patent include its speed, costs, and efficiency. The approach is very sensitive to the secondary structure of oligonucleotides (e.g., hairpin loops) which interfere with the assembly. Hence, the approach has low efficiency and is not reliable for construction of long synthetic DNA fragments. Moreover, the method involves numerous steps and requires synthesis of both strands of the DNA. It requires purification of many short (50-70 nt) oligonucleotides by rather laborious and time consuming procedures. Use of T4 DNA ligase requires conventional temperatures (20° – 37°C) during assembly and ligation of oligonucleotides therefore contributing to the sensitivity of the method to secondary DNA structure.

Another approach to gene synthesis is described in U.S. Patent No. 4,652,639 to Stabinsky. That reference also relates to a method for DNA synthesis wherein the synthetic gene has alternative codons selected on the basis of preferential expression in a projected host organism to be transformed. The method employed involves the ligation of two or more DNA strands although no template is employed. A key disadvantage of the gene synthesis method described in the '639 patent is that it can only be employed to synthesize short genes of about 200 base pairs. It is, however, frequently desirable to create longer genes.

Yet another method for the synthetic assembly of oligonucleotides into long DNA fragments utilizes polymerase to fill in single-stranded gaps in annealed pairs of oligonucleotides. However, after the polymerase reaction, each segment must be cloned, a step which significantly delays the synthesis of long DNA fragments and greatly decreases the efficiency of the approach. Additionally, the approach can be used only for small DNA fragments.

Recently, several PCR based techniques have been developed for construction of synthetic genes, where an assembly of overlapping oligonucleotides is performed by a thermostable DNA polymerase during repeated cycles by melting, anneling and polymerization. Although PCR mediated methods are rather simple and labor saving, they are not free from drawbacks. A key disadvantage of assembly by PCR is that complex mispriming events negatively exert the correctness of a resulting assembled DNA. In addition, the low fidelity of thermostable DNA polymerase influences the reliability of this technology with increased number of PCR steps.

Template Directed Ligation

There is known in the art a ligase mediated method for detecting mutations that exploits the fact that the ends of two single strands of DNA must be exactly aligned for DNA ligase to join them. U.S. Patent No. 5,547,843 to Studier et al, briefly summarizes a template-directed ligation method wherein two contiguously annealing oligonucleotides are incubated with a template molecule which contains a nucleotide sequence complimentary to the oligonucleotide. The oligonucleotides anneal to form a contiguous duplex structure. The complex is then contacted with a ligase enzyme which joins the adjacent oligonucleotide through a phosphodiester linkage. The only utility suggested in the '843 patent and the prior art in general for template-directed ligation is as a diagnostic method for the detection of point mutations in DNA. Heretofore, there has been no teaching or suggestion in the art of utilizing template directed ligation for the assembly and construction of synthetic genes in a manner which overcomes the above-described disadvantages of prior art techniques.

SUMMARY OF INVENTION

The present invention is a new method for assembly and construction of synthetic genes which are similar but not necessarily identical to a natural gene of interest. The method utilizes a combination of enzymatic and chemical synthesis of DNA and significantly reduces the cost, time and number of steps required for construction of synthetic genes.

In short, chemically synthesized and phosphorylated oligonucleotides of the gene to be created are assembled on a single-stranded partially homologous template DNA derived from the natural or wild-type gene. After annealing, the nicks between adjacent oligonucleotides are closed by a thermostable DNA ligase using repeated cycles of melting, annealing, and ligation. This template directed ligation ("TDL") results in a new single-stranded synthetic DNA product which is subsequently amplified and isolated from the wild type template strand by the polymerase chain reaction (PCR) with short flanking primers that are complementary only to the new synthetic strand. These PCR end-primers contain suitable restriction cleavage sites for cloning of the synthetic double-stranded DNA fragments.

There are many advantages of the "TDL-PCR" gene synthesis method according to the present invention over the above-described prior art approaches. First, only one strand is required to be synthesized thereby reducing the time requirements by about one half. Second, the use of thermostable DNA ligase permits the process to be carried out at higher temperatures which decreases the formation of secondary structures (e.g. hairpin loops) that interfere with the assembly process. Third, because only phosphorylated oligonucleotides will be ligated during repeated TDL cycles there is no need to purify full length oligonucleotides. Fourth, the process involves fewer steps than the prior art approaches thereby increasing efficiency and decreasing the likelihood of errors.

In the TDL-PCR method the PCR is used only in the final step, thus avoiding the assembly of oligonucleotides, in contrast to PCR based methods for gene synthesis described in the preceding section. The gene assembly by ligation of oligonucleotides is preferable to assembly by PCR approach, because no new errors are introduced to the assembled DNA. Indeed, the analysis of the observed errors showed (Example 1) that these originate from chemical oligonucleotide synthesis - a common problem with any gene synthesis technology, and that using the PCR in the final step did not introduce additional errors.

The disclosed TDL-PCR method for gene synthesis represents a general method which may be employed to create a synthetic version of a natural or wild-type gene from any plant or animal or any existing gene as long as the wild-type sequence is known and present. The simplicity of the method allows it to be widely used in areas where the routine re-synthesis of the sequences is required. There is an increasing demand for processes that alter DNA sequences to permit the construction of transgenic organisms. As a result of billions of years of evolution, microorganisms have genes in their genome that are not present in other higher organisms such as plants and vertebrates. Frequently such genes are not suited for expression in eukaryotic organisms and vice- versa. Moreover, even in plants the monocot and the dicot genes differ by expression requirements, and therefore need to be reconstructed to fit different expression patterns.

With the foregoing and other objects, advantages and features of the invention that will become hereinafter apparent, the nature of the invention may be more clearly understood by

reference to the following detailed description of the invention, the figures, and the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a schematic plan for gene synthesis according to the present invention.

Fig. 2 shows the nucleotide sequence of the synthetic *cryIC* gene. Where different, the native (bacterial) sequence is shown above. The amino acid sequence is shown below.

DETAILED DESCRIPTION OF THE INVENTION

Prior to initiating the synthesis process, the sequence of the new gene to be created is designed based on the known sequence of the wild-type gene. If the new gene will be used to transform a plant, then modifications will be made with respect to the plant's preferred codon's and to other fortuitous sequences resembling the signals which initiate various processing steps in transcription and translation of the gene. Avoidance of such signals can drastically increase expression of a transgenic in plants.

Gene synthesis according to the present invention generally comprises three steps which are illustrated schematically in Fig.1.

During the first step a plurality (e.g. eight) of oligonucleotides, each of a length of between 80-130 bases which collectively comprise the newly designed sequence, are chemically synthesized and phosphorylated by the conventional solid-phase method. A chemical phosphorylation is performed during the last step of automated oligonucleotide synthesis so that only the full- length oligonucleotides have a 5'-phosphate group. By this arrangement only phosphorylated (i.e. full-length) oligonucleotides can be ligated with other adjacent oligonucleotides during the following TDL step thereby conveniently eliminating oligonucleotides that are too short.

The second step known as template directed ligation (TDL) results in a new synthetic single-stranded DNA product, which is subsequently converted to a double-stranded fragment, amplified and selected from the wild type strand in the following third step by PCR selective amplification comprises the following procedures:

(a) Assembly of the adjacent chemically synthesized and phosphorylated oligonucleotides on the single-stranded partially homologous DNA template of the natural (wild type) gene allowing the oligonucleotides to anneal to the template in their proper order;

(b) Selection of the full-length oligonucleotides from a crude mixture of oligonucleotides after chemical DNA synthesis. Due to chemical phosphorylation only full-length oligonucleotides can be ligated with other assembled adjacent oligonucleotides, and therefore they are selected from a crude mixture. In case of an average oligonucleotide length between 100 and 120 nucleotides with a coupling efficiency around 98.0 - 98.5%, what is typically used, only 10-20% of oligonucleotides in crude mixture constitute the full- length DNA products which are included in the synthesis method. In the prior art, the full-length oligonucleotides must be purified from the crude mixture by laborious methods. However, in the TDL-PCR method this is not necessary, because uncompleted oligonucleotides cannot be ligated and , therefore, they are not amplified during the following PCR step. This selection of full-length oligonucleotides makes the TDL-PCR method according to the present invention considerably more advantageous over other known methods for a gene synthesis;

(c) Nick closing. Adjacent oligonucleotides assembled on the single-stranded wild- type DNA template are ligated together thereby closing nicks between adjacent oligonucleotides. Partially ligated DNA fragments can be completed in the following repeated rounds of the TDL step. Contrary to an exponential amplification in the ligase chain reaction (LCR), the amount of the final product is increased linearly during cycling in TDL step, if there is a sufficient excess of free oligonucleotides. Therefore, the yield depends on the ratio of template to oligonucleotides. The *Pfu* DNA ligase employed is extremely thermostable and possesses high nick-closing activity in contrast with low level of blunt ended (template-independent) ligation activity and therefore was used and is recommended for this application.

During the third and final step of the TDL-PCR method according to the present invention the newly assembled and sealed single-stranded DNA is converted to the double-stranded fragment and subsequently amplified and selected from the wild type template strand by the PCR with short flanking primers that are complementary only to the new synthetic strand.

These flanking primers also contain sites for selected restriction enzymes in order to facilitate the cloning of an amplified DNA fragment.

The following example is provided to illustrate the practice of the invention and is not intended to limit the scope thereof.

EXAMPLE I

Synthesis of *Bt* toxin gene

In this example, the above-described TDL-PCR method was employed to synthesize a *cryIC* gene which encodes a *Bacillus thuringiensis (Bt)* endotoxin. The natural gene was found to be poorly expressed in transgenic plants. Efficient transcription of native *cry* genes in plant cell nuclei was achieved by the removal of AT-rich sequences that cause mRNA instability or aberrant splicing, and the translation of *cry* mRNAs was enhanced by modification of their codon usage to make it more similar to that of the host plant. In addition, the sequence context around translation start was modified to conform the eukaryot consensus (Kozak, 1990).

Design of the Synthetic Gene

Design of the synthetic *cryIC* gene was based on the sequence of the corresponding wild-type gene (*cryICa5*, as in Fig. 2). In fact, the *cryICa5* sequence represents a consensus of all known *cryIC* genes and is identical in three *B.thuringiensis* strains K26-21, MR1-37 and sub sp.*aizawai* 7.29. Modifications of the synthetic *cryIC* gene (*s-cryIC*) sequence did not alter the amino acid sequence of the minimal toxic fragment of the Cry IC protoxin, containing N-terminal fragment with length of 630 amino acid residues (Fig 2). To ensure proper transcription and translation of this synthetic gene in plants, 258 bp of the bacterial *cryIC* sequence (EMBL X96682; 1890 bp) were changed such that 249 out of 630 codons were modified according to preferential codon usage in alfalfa and other dicotyledonous plants. These nucleotide exchanges also removed 21 potential plant polyadenylation signals, 12 ATTTA motives, 68 sequence blocks with 6 or more consecutive A or T nucleotides, as well as all sequence motifs containing 5 or more G+C or A+T nucleotides. In addition, sequences around the translation initiation site were changed in order to conform the eukaryotic consensus sequence, and a TAG stop codon

was introduced downstream of amino acid codon 630. The G+C content of the bacterial *cryIC* gene, was thus increased from 36.6% to 44.8% in the synthetic gene.

Oligonucleotide Synthesis

Oligonucleotides were synthesized with an Applied Biosystems 380B DNA synthesizer, using a 1000Å controlled pore glass support (Millipore Eschborn, Germany). Except for primers located at the 5'-termini of TDL-PCR sequence blocks (see below), the 5'-ends of oligonucleotides were chemically phosphorylated with 5'-Phosphate-ONphosphoramidite (Clontech, Heidelberg, Germany) according to manufacture's manual in the final step of DNA synthesis. The synthesis products were deprotected with ammonium hydroxide, desalted on NAP25 (Pharmacia,Freiburg, Germany) columns, and used for TDL-PCR without additional purification.

Gene Construction

The designed DNA sequence of the *s-cryIC* gene (Fig. 2) was divided into three blocks separated by *Hinc*II and *Bgl*II cleavage sites. The *BamHI-Hinc*II block-I was constructed from eight, the *Hinc*II - *Bgl*II block-II from five, and the *Bgl*II-*Bam*HI block-III from seven oligonucleotides. The oligonucleotides were assembled on a single-stranded DNA template of phagemid pR1, carrying the 630 N-terminal codons of the wild-type *B.thuringiensis cryIC* gene (Fig. 1 and 2). Terminal oligonucleotides in each TDL-PCR block carried unique sequences on their 5' and 3' ends, which were not complementary with the template, but were matched to short PCR primers for selective amplification of the synthetic DNA strand. These PCR primers contained unique restriction enzyme cleavage sites used for cloning of the amplified double-stranded DNA fragments into pBluescript. The TDL-PCR block-I was PCR amplified by a 5'-primer (5' -AAGAGGATCCACCATGGAGGAGAAC-3 '), carrying a *Bam*HI site and a 3'-primer (5'-ATGATCTAGATGCAGTAGCG-3'). The 3'-primer was complementary to an oligonucleotide (5 '-GTCAACTAACAAGGGAAGTTTATACGGACCCACGCT ACTGCATCTAGATCAT-3') at the 3'-end of block-I, that carried *cryIC* sequences with the *Hinc*II site, and unrelated overhang sequences with an *Xba*I site. The oligonucleotide at the 5'-end of block-II(5'-GATAACTCGAGCGAGCCTAAACTATGACAATAGGAGATATCCAATTCAG

CCAGTTG - 3') added unique DNA sequences with an *Xho*I site to the *cryIC* sequences upstream of the *Hinc*II site and matched a PCR primer (5'-GATAACTCGAGCGAGCCTA-3'). The 3'-terminal oligonucleotide in block-II carried *cryIC* sequences extending to the *Bgl*II site and downstream overhang sequences with an *Xba*I site that were complementary to a PCR primer (5' -CCTGACTCTAGAAGATC-3 '). In the oligonucleotide located at the 5'-end of block-III an *EcoR*I site was added upstream to the *Bgl*II site of *cryIC* gene, fitting to a PCR primer (5'- CTGTCTGAATTCAAAGATC- 3'). The oligonucleotide at the 3'-end of block-III carried a *BamH*I site, following the position of TAG stop codon in the pR1 phagemid, as well as adjacent unique sequences with a NotI site that were complementary to a PCR primer (5'-AGCATGCGGCCGCGGATCC-3').

TDL Technique

Template directed ligation (TDL) reactions were carried out at a template to oligonucleotide ratio 1:200 (a total of 0.05 pM of template versus 10 pM of each oligonucleotide) in a final volume of 50µl using a reaction buffer (20 mM Tris-HCl (pH 7.5), 20 mM KCl, 10 mM $MgCl_2$ 0,1% NP-40, 0.5 mM rATP, 1 mM DTT)-and 4 U *Pfu* DNA ligase (Stratagene), or any other similar thermostable DNA ligase.

Thirty cycles of TDL reactions were used to obtain a desirable amount of a TDL product. The temperature range during melting step is between 90 to 98°C with a preferable temperature of 92°, with 1 minute of required step time. Annealing and ligation were performed at a temperature range of 45 to 60°C with a preferable temperature of 52°C during required step time from 3 to 10 minutes. Melting step was followed by annealing and ligation step to obtain a TDL cycle which was repeated at least 30 times. To increase the number of TDL cycles for every additional 30 cycles a new portion of rATP (0.5 mM) and 4 U of *Pfu* ligase was added. Temperature cycling during TDL step was done on a Perkin-Elmer thermal cycler (Norwalk, Conn.).

PCR Selective Amplification of Synthetic TDL-PCR Blocks

5 µl from the TDL reaction mix served as template for PCR amplification with 100 pM of primers, 250 µM dNTP and 2.5 U Ampli-Taq or any other similar thermostable DNA

polymerase such as UlTma (Perkin-Elmer) polymerase in 100 µl buffer (10 mM Tris.HCl (pH 9.0), 50 mM KCl and 0.1% Triton- X100), using of 30 cycles at 92°C for 1 min, 45°C for 1 min, and at 72°C for 1.5 minutes, with final extension for 10 minutes, at 72° C PCR amplifications were performed on a Perkin-Elmer (Norwalk, Conn.) thermal cycler. The amplified DNA fragments were gel purified, digested with *Bam*HI-*Xba*I (block-I), *Xho*I-*Xba*I (block-II), and *Eco*RI-*Not*I (block-III), then cloned in pBluescript SK+ to verify their DNA sequences.

The synthetic *cryIC* gene coding for an N-terminal protoxin fragment of 630 amino acids was designed (Fig. 2) by exchanging 286 bp of the bacterial *cryIC* sequence (EMBL X96682; 1890 bp) such that 249 out of 630 codons were modified according to preferential codon usage in dicotyledonous plants. These exchanges removed 21 potential plant polyadenylation signals, 12 ATTTA motifs, 68 sequence blocks with 6 or more consecutive A/T's, and all motifs containing 5 or more G+C or A+T nucleotides. Sequences around the translation initiation site were changed to conform to the eukaryotic consensus sequence, and a TAG stop codon was introduced downstream of amino acid codon 630. The G+C content of the *cryIC* gene was thus increased from 36.6% to 44.8%. The *s-cryIC* gene was synthesized from oligonucleotides of 70-130 bases that were chemically phosphorylated at their 5'-ends. Since chemical phosphorylation is performed as the last step of automated DNA synthesis, only full-length oligonucleotides contain the 5'-phosphate group. Bacterial *cryIC* sequences coding for the 630 N-terminal codons were cloned in a pBluescript vector to generate a single- stranded DNA template for ordered annealing of 5-8 synthetic oligonucleotides by partial base-pairing. The adjacent oligonucleotides were assembled and ligated on this single- stranded template by a thermostable *Pfu*-ligase using 30-60 cycles of repeated melting, annealing and ligation. In combination with chemical phosphorylation this template directed ligation (TDL, Fig. 1) method provided a sequence specific selection for phosphorylated full-length oligonucleotides from a complex mixture of nonphosphorylated failure synthesis products, and yielded a linear amplification of single-stranded synthetic *cryIC* DNA segments generated by ligation. Therefore, except for desalting, no additional purifications of a crude oligonucleotide mixture after chemical DNA synthesis were necessary. The TDL ligation at high temperatures also circumvented potential problems of erroneous annealing. The synthetic *cryIC* sequences were converted to double-stranded DNA fragments and specifically amplified by PCR using short end-primers that did not anneal to the bacterial *cryIC* template carried by the pBluescript vector. The *s-cryIC* gene

was thus synthesized from three sequence blocks that were combined by ligation of *Hinc*II and *Bgl*II digested DNA fragments, and cloned in pBluescript.

With further reference to the Figures relating to this Example, Fig. 2 shows the nucleotide sequence of the synthetic *cryIC* gene *(s-cryIC)*. Nucleotides of the bacterial *cryIC* sequence (*b-cryIC*) exchanged in the synthetic gene are shown in the upper lanes. The amino acid sequence of the truncated CryIC- endotoxin is displayed in single letter code below the *s-cryIC* sequence. The nucleotide sequence of the *s-cryIC* region coding for 630 codons starts with an ATG codon in a sequence context fitting the eukaryotic consensus and terminates at aTAG stop codon. Vertical black arrows above the *s-cryIC* sequence indicate the boundaries of adjacent synthetic oligonucleotides used for TDL-PCR gene synthesis. *Hinc*II and *Bgl*II cleavage sites used for the assembly of three TDL-PCR blocks are framed.

The errors (3 small deletions, 1 transversion and 1 transition) found in the TDL-PCR products were corrected by site-directed mutagenesis using a USE kit (Pharmacia) or by assembly of non-mutated restriction fragments.

Plant gene expression constructs and transformation of alfalfa and tobacco.

The plant expression vector pPCV91 was constructed by modification of pPCV720. A *Not*I site in the RK2-domain was eliminated by filling in with DNA polymerase Klenow fragment, and a CaMV35S promoter with four repeats of the enhanced domain (-90 to - 418), was introduced into the *Hind*III site of pPCV720. Upstream of a *Bam*HI cloning site this cassette contained 20 bp from the 3'-end of the untranslated Ω leader sequence of tobacco mosaic virus (TMV) RNA, whereas downstream of the *Bam*HI site it carried a polyadenylation signal sequence derived from the CaMV 35S RNA gene. A *Bam*HI site present in the mannopine synthase promoter(pmas) of pPCV720 was replaced by a *Not*I site using a *Sau*3A-*Not*I adaptor (5'-GATCTGCGGCCGCA-3'). The resulting vector pPCV91 carried three plant gene expression cassettes with unique *Bam*HI, *Not*I and *Sal*I cloning sites. To construct pNS6, the synthetic *cryIC* gene was cloned as a *Bam*HI fragment downstream of the CaMV35S promoter. In pNS7 a synthetic *pat* gene, coding for a phosphinothricine acetyltransferase (*pat*) gene and a *chi*AII gene from *Serratia marcescens* were inserted into the *Sal*I and *Not*I sites located respectively downstream

of the *mas* 1' and 2' promoters. A bacterial *cryIC* gene from B.*thuringinesis* sub sp.. *aizawai* 7.29 (EMBL X96682), carrying the 756 N-terminal codons of *cryIC*, was cloned in pGIF1 in which it replaced the synthetic *cryIC* gene of pNS7. Vectors pNS6, pNS7 and pGIF1 were conjugated to *Agrobacterium tumefacines* GV3101(pMP90RK), and used for transformation of alfalfa (*Medicago sativa* L. var. Regen S clone RA3) and tobacco (*Nicotaina tabacum* SRI) as described. To select for transformed explant, alfalfa and tobacco tissue culture media contained respectively 40 µg /ml and 15 µg/ml of hygromycin were used.

Monitoring the expression of CryIC in transgenic plants

Bacterial and synthetic *cryIC* genes, coding for the 630 N-terminal amino acids of the CryIC toxin (Fig. 2), were cloned into the *Bam*HI site of a pAEN4 vector carrying the CaMV35S gene expression cassette of pPCV91. *Arabidopsis thaliana* protoplast were isolated from root cultures and transformed by PEG-mediated DNA uptake, using 1.5 x 10^6 protoplast and 35 µg plasmid DNA in each experiment. The protoplast were harvested 48 hours after DNA uptake and lysed in SDS-sample buffer to separate proteins on 10% SDS-PAGE before immunoblotting. An antibody used for immunoblotting was raised against a truncated CryIC δ-endotoxin carrying 756 N-terminal amino acids. Expression of CryIC in *E.coli* strains, carrying bacterial or synthetic *cryIC* genes respectively in pET-lla or 11d, was monitored-by a second alkaline phosphatase conjugated goat anti-rabbit antibody. Immunoblott analysis of proteins synthesized in plant cells was performed using an ECL kit (Amersham).

RNA (20 µg) samples isolated from leaves and petioles of alfalfa plants were separated on agarose-formaldehyde gels. *Bam*HI fragments (1.9 kb), carrying either synthetic or bacterial *cryIC* sequences (Fig.2), and a *Not*I fragment of the *chiAII* gene (1.8 kb) were labeled by random-priming and used as hybridization probes.

Insect bioassay

Leaf bioassay were performed with the Egyptian cotton leafworm (*Spodoptera littoralis*) and the beet armyworm (*Spodoptera exigua*) using neonate, 2-3rd. 3-4th, and 4-5-6th instar larvae. Ten larvae of selected developmental stage were placed on a moistened filter disc in Petri

dishes with detached leaves from greenhouse grown plants. The assays were repeated 2-3 times for each plant. The mortality of neonate larvae was scored after 3 days, whereas the mortality of larvae from 2-4th and from 4-6th instar stages were evaluated respectively after 5 and 7 days. For the insect assays with whole plants, transgenic greenhouse grown alfalfa lines producing 0.02 to 0. 1% of total soluble protein as CryIC and S.exigua larvae of the 3-4th instar stage were used. Three NS7 and three NS6 transgenic, as well as wild-type plants were infested with 15-20 larvae each. In "free-choice" experiments, 25 larvae were placed in a Petri dish located between transgenic NS6 or NS7 and nontransgenic alfalfa plants in the greenhouse. Leaf damage was evaluated after 6 days.

Further details regarding this example may be found in our co-pending application titled "Synthetic *Bacillus Thuringiensis* Gene Encoding an Insect Toxin".

Although only preferred embodiments are specifically described herein, it will be appreciated that many modifications and variations of the present invention are possible in light of the above teachings and within the purview of the appended claims without departing from the spirit and intended scope of the invention.

What is claimed is:

1. A method for constructing a synthetic version of a native or any existing gene comprising the steps of:

providing a DNA strand of the native gene;

denaturing said DNA strand to form a single-stranded template;

chemically synthesizing a plurality of oligonucleotides at least partially homologous to said single-stranded template;

phosphorylating said oligonucleotides;

annealing said oligonucleotides to said single-stranded template;

ligating adjacent said oligonucleotides to form a synthetic gene; and,

repeating said steps of denaturing, annealing, and ligation to form a useful quantity of said synthetic gene.

2. The method according to claim 1, further comprising the steps of converting said synthetic gene sequence to double-stranded DNA and amplifying said double-stranded DNA using PCR.

3. The method according to claim 2, wherein short flanking primers complementary only to the new synthetic strand are utilized in PCR.

4. The method according to claim 3, wherein said flanking primers further comprise sites for selected restriction enzymes so as to facilitate cloning of the amplified DNA fragment.

5. The method according to claim 1, wherein said ligation is carried out by a thermostable DNA ligase.

6. The method according to claim 5, wherein said thermostable DNA ligase is *Pfu* DNA ligase.

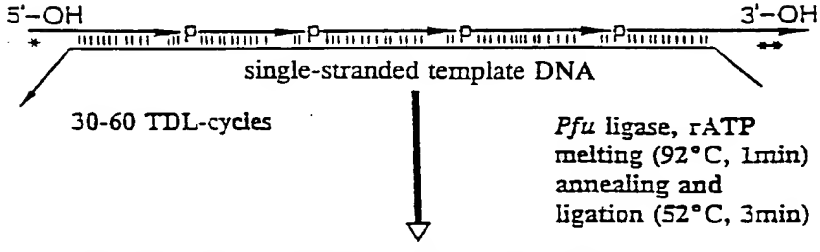
7. The method according to claim 1, wherein said native DNA strand is a *Bacillus thuringiensis* (*Bt*) toxin gene.

8. The method according to claim 6, wherein said *Bacillus thuringiensis* (*Bt*) toxin gene is a *cryIC* gene.

9. The method according to claim 1, wherein said phosphorylating is a phosphorilation performed at the last step of automated oligonucleotide synthesis.

1: Automated synthesis of 5'-phosphorylated oligonucleotides

2: Template directed ligation of oligonucleotides annealed to a partially complementary single-stranded DNA carrying wildtype gene sequence



3: Selective PCR amplification of synthetic DNA strand

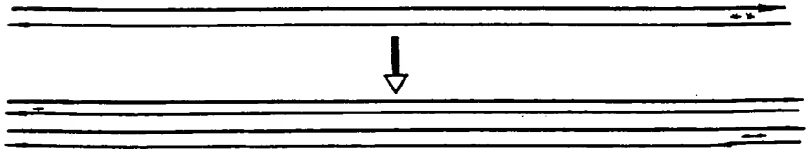


Fig. 1

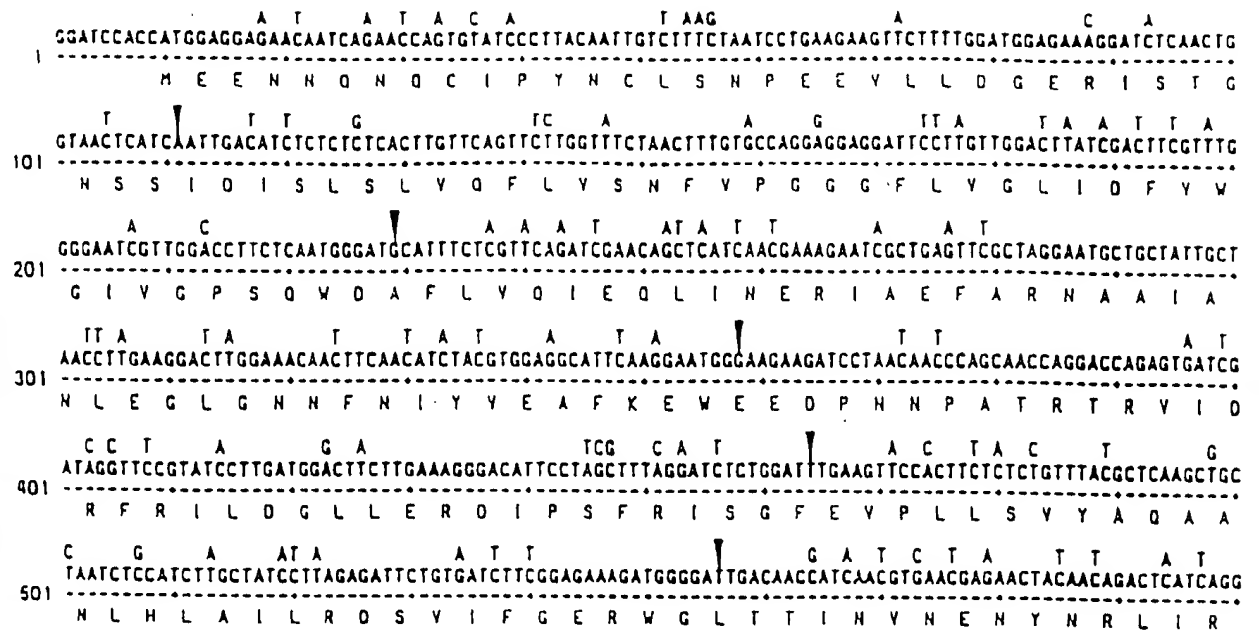


Fig. 2

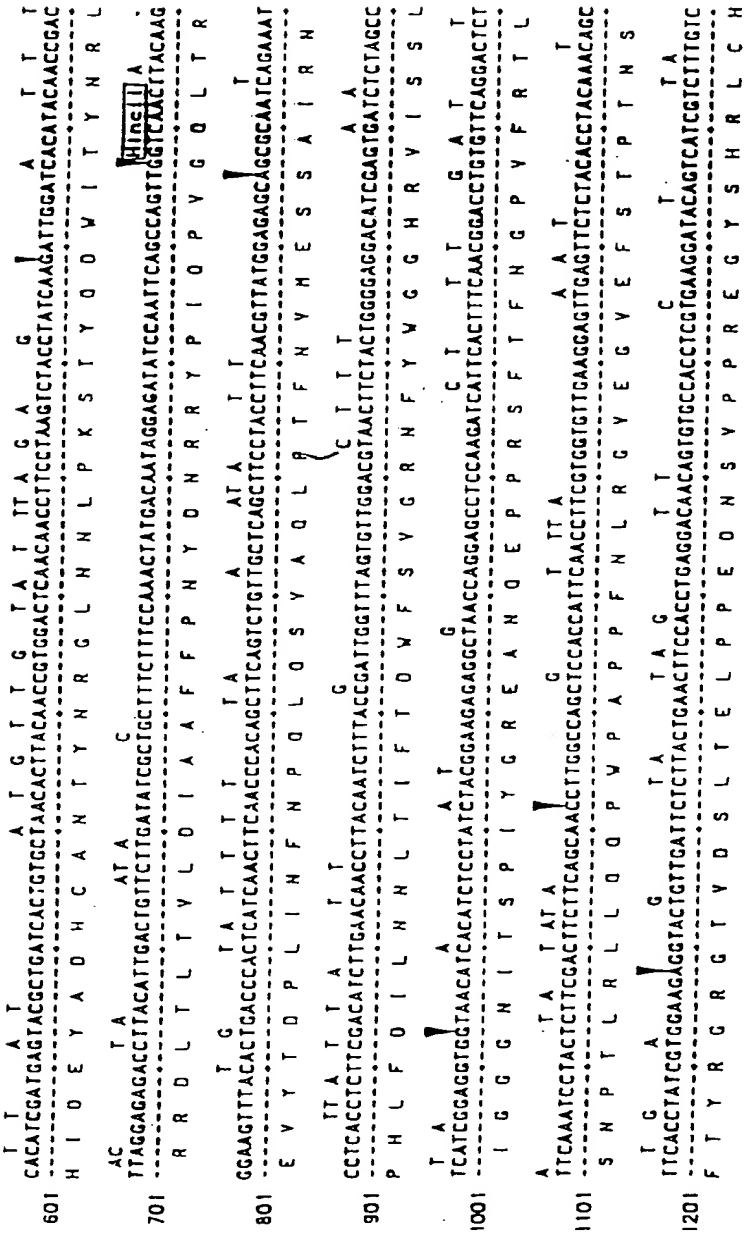


Fig. 2 (cont'd)

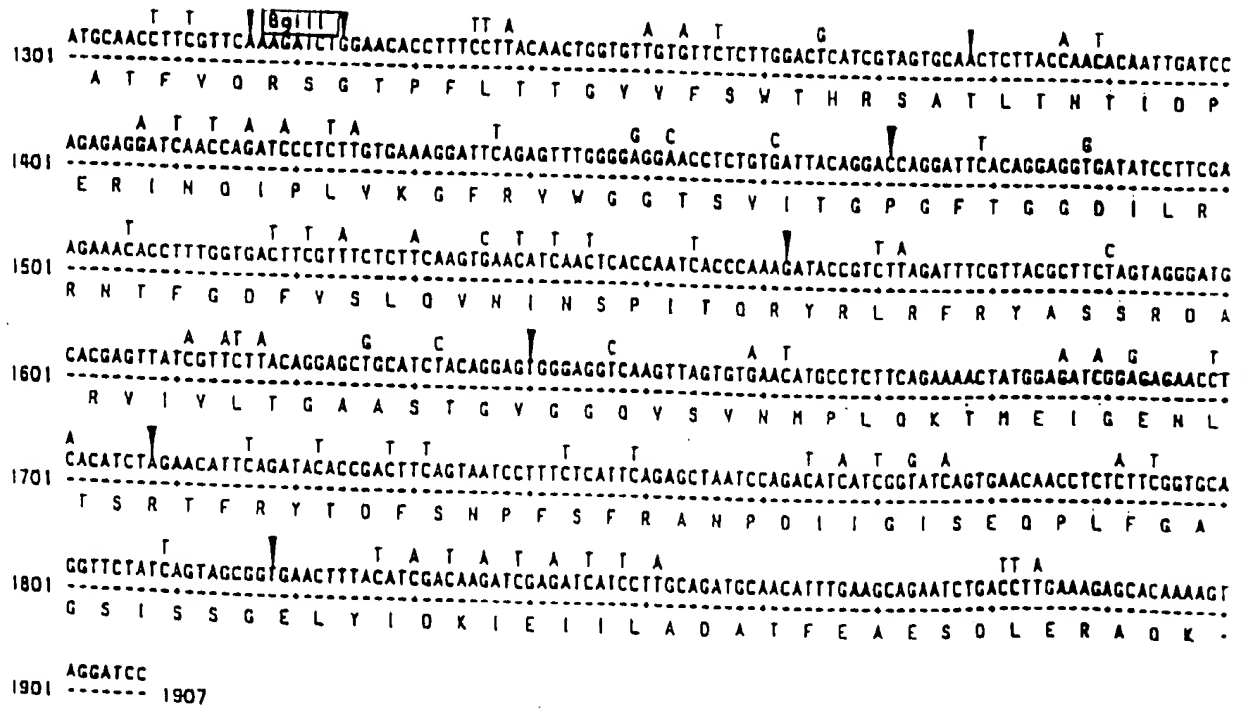

1301
ATGCAACCTTCGTTCAAAGATCTGGAACACCTTTCCTTACAACTGGTGTTGTGTTCTCTTGGACTCATCGTAGTGCAACTCTTACCAACACAATTGATCC
A T F V Q R S G T P F L T T G Y V F S W T H R S A T L T N T I D P
1401
AGAGAGGATCAACCAGATCCCTCTTGTGAAAGGATTCAGAGTTTGGGGAGGAACCTCTGTGATTACAGGACCAGGATTCACAGGAGGTGATATCCTTCGA
E R I N Q I P L V K G F R V W G G T S V I T G P G F T G G D I L R
1501
AGAAACACCTTTGGTGACTTCGTTTCTCTTCAAGTGAACATCAACTCACCAATCACCCAAAGATACCGTCTTAGATTTCGTTACGCTTCTAGTAGGGATG
R N T F G D F V S L Q V N I N S P I T Q R Y R L R F R Y A S S R D A
1601
CACGAGTTATCGTTCTTACAGGAGCTGCATCTACAGGAGTGGGAGGTCAAGTTAGTGTGAACATGCCTCTTCAGAAAACTATGGAGATCGGACAGAACCT
R V I V L T G A A S T G V G G Q V S V N M P L Q K T M E I G E N L
1701
CACATCTAGAACATTCAGATACACCGACTTCAGTAATCCTTTCTCATTCAGAGCTAATCCAGACATCATCGGTATCAGTGAACAACCTCTCTTCGGTGCA
T S R T F R Y T D F S N P F S F R A N P D I I G I S E Q P L F G A
1801
GGTTCTATCAGTAGCGGTGAACTTTACATCGACAAGATCGAGATCATCCTTGCAGATGCAACATTTGAAGCAGAATCTGACCTTGAAAGAGCACAAAAGT
G S I S S G E L Y I D K I E I I L A D A T F E A E S D L E R A Q K .
1901
AGGATCC
------- 1907

Fig. 2 (cont'd)